DOCUMENT RESUME

ED 111 848 TM 004 815

AUTHOR	Baker, Robert L.
TITLE	Measurement Considerations in Instructional Product Development. Professional Paper No. 29.
INSTITUTION	Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO	SWRL-PP-29
PUB DATE	Mar 74
NOTE	20p.
EDRS PRICE	MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS	Criterion Referenced Tests; Curriculum Development; *Educational Development; Educational Objectives; *Instructional Improvement; Instructional Materials; *Instructional Programs; Instructional Systems; Measurement; *Program Development; *Psychometrics; Test Construction; Testing Problems

ABSTRACT

The application of conventional psychometric procedures to instructional product development is outlined. Selected non-psychometric requirements of developing and delivering effective instruction are described and the consequent implications for psychometric procedures set forth. Present instructional development technology is circumscribed in terms of state-of-the-art capability. (Author)

SWRL

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Measurement Considerations In Instructional Product Development

20 March 1974, Professional Paper 29

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Professional Paper 29

March 1974

MEASUREMENT CONSIDERATIONS IN INSTRUCTIONAL PRODUCT DEVELOPMENT

Robert L. Baker

ABSTRACT

The application of conventional psychometric procedures to instructional product development is outlined. Selected non-psychometric requirements of developing and delivering effective instruction are described and the consequent implications for psychometric procedures set forth. Present instructional development technology is circumscribed in terms of state-of-the-art capability.

MEASUREMENT CONSIDERATIONS IN INSTRUCTIONAL PRODUCT DEVELOPMENT*

Robert L. Baker

The psychometric revolution that has been smoldering over the past decade and finally ignited in the "criterion-referenced test movement" will predictably spread throughout education during the next decade, and will generate consequences that go well beyond the boundaries of psychometry (Schutz, 1972). Even now it is obvious that concern with psychometric dogma reflected in such questions as "Is the criterion-referenced test just a special instance of the norm-referenced test?" and "How can the reliability of criterion-referenced tests be assessed?" is misplaced. Focusing on such questions is about as productive as the programmed instruction research of the 1960's related to overt-covert and large step-small step issues.

Recent instructional research and development has demonstrated that formal measurement can indeed fulfill important roles in producing instructional programs to meet prespecified objectives. However, full exploitation of this role requires control over non-psychometric as well as psychometric variables. The purely technical aspects of psychometry provide great capability for instructional product development. Conventional psychometric procedures can readily be adapted to generate measures which provide adequate bases for those instructional decisions that can currently be made and effected. But this is inadequate to advance the state-of-the-art for improving instructional effectiveness.

*Originally prepared as part of the symposium "Evaluation of Instructional Systems: Current Problems Facing Developers," American Psychological Association Annual Meeting, Honolulu, Hawaii, September 3, 1972.

The interface between psychometry and instructional development must include greater attention to instructional decision algorithms that are defined as functions of achievement measures anchored systematically to the manipulable conditions that produced the achievement. This considera-tion will encompass not only the specifications and development of instruction but also the installation and continuing operation of instruction. The effectiveness of specified instructional decision algorithms is dependent upon well-defined assessment procedures that are easily reflected in defined behavioral classes of interest and anchored in manipulable instructional determinants.

The manipulable determinants of achievement in developing instructional programs are materials and procedures. To be useful in a development context tests must be designed and constructed in a manner that defines the explicit rules linking patterns of test performance to behavioral referents anchored in sequenced instructional materials and procedures. Further, to be useful in an operating instructional context tests must be configured in such a way that a particular decision algorithm may be applied with little inconvenience.

The testing requirements following from these conditions are manifold, and the scientific and technological bases for getting on with it range from adequate to non-existent. However, absence of these bases cannot halt development efforts. We must identify the immediately available resources for developing effective instruction and move as quickly as possible to completion of first generation shelf items, recognizing that the items thus produced represent only a beginning of "more to come" from programmatic educational R&D currently in progress.

The remainder of the paper will view selected psychometric requirements and strategies as they interface with selected non-psychometric requirements of developing and delivering effective instruction. The view will be from within defined SWRL R&D activities and state-of-the-art capability.

MATTERS THAT ARE WELL WITHIN SWRL STATE-OF-THE-ART

The instructional development technology described in this section is readily available in shelf-item or easily adaptable form.

Writing Instructional Objectives

The "how-to" information for stating well-formed instructional objectives has been available for some time. A convenient recent synthesis of this information is contained in the SWRL Staff Development Compendium (Baker & Schutz, 1971). By reading this information, an interested high school-graduate-equivalent person can acquire all of the information required to meet this condition. The time-consuming and thought-challenging task of what outcomes to prepare remains to be done. But this is a matter of doing the job, rather than of not knowing how.

When the job of preparing well-formed instructional outcomes has been completed, one is at best at the beginning rather than at the end of instructional effectiveness. But the beginning is firm, rather than wishful.

Criterion-Referenced Test Construction

Not only does instructional development necessitate prespecified instructional outcomes, it also requires a means of assessing the attainment of these outcomes. This involves test construction activity.

To be minimally useful the tests must be specifically referenced to a prespecified structure of achievement. To be maximally useful the tests must be specifically referenced to defined instructional materials. A consequent requirement is to define criterion behavior in the specification of the limits of a population of responses called for in the instruction which defines the criterion behavior rather than in a list of responses which exemplify it. This is not a new concept; it was encompassed by earlier discussions of content validity (Lennon, 1956) and Bruner's (1960) discussion of the structure of the subject matter knowledge. However, specific procedural cues for meeting the requirement were not available until Hively (1963) introduced the "item form."

The item form and related processes provide a neat system for blueprinting tests that meet all of the requirements of the psychometric concept of content validity and at the same time contribute to the definition of the behavioral structure of the subject matter domain treated. A collection of item forms sequentially ordered, together with the replacement sets for the variable elements, could adequately define a universe of content across specified outcome areas. When such procedures are more generally exploited the impracticality of constructing criterion-referenced tests for complex behavioral-content domains cited by Ebel (1971) is overcome.

Instructional Specifications

An item form defines classes of behavior, but it does not indicate how the behavior is to be established. However, as strings of item forms are prepared, it is possible to arrange them into tentative

sequences that constitute an operational "cognitive map" of a subject matter useful in guiding both instructional and evaluational efforts.

The "instructional specifications" approach (Sullivan, Baker & Schutz, 1971) provides a set of procedures for maping out the instructional and assessment sequences consistent with the item form. The instructional specification (IS) is a convenient guide to the development of effective instruction for a given instructional objective. A well constructed IS per instructional objective provides answers to the following questions:

1. What outcomes (objectives) will the successful learner attain as a result of the instruction?
2. What information (cue) will be given the learner to increase his ability to perform the desired behavior?
3. What procedures (mastery items) will be used to provide for practice and assessment of the desired behavior?
4. What are the characteristics (limits) of the correct responses or response choices for the desired behavior and what are the characteristics of plausable but incorrect responses?
5. What relevant skills (entry skills) must the learner possess prior to the instruction for the present objective?

Instructional programs that are developed properly from a set of written IS's incorporate the instructional and assessment techniques directly into the program materials and procedures, thereby increasing the probability of high learner achievement of the instructional objectives.

The IS is primarily useful in specifying instruction prior to the development of materials and procedures. However, the structure and architecture of extant instruction and curricula are seldom explicitly stated. Postdictive analytic conventions (Smith, 1971) have been developed for use in analyzing the instructional architecture of portions of instructional materials. Set and matrix notational conventions permit description of extant material in terms of the following seven components:

<u>Elements</u>: the phenomena to be described, compared, related, or otherwise studied (e.g., objects, systems, events, groups).

<u>Variables</u>: the characteristics of properties of elements that are used to describe, compare and relate them (e.g., color, weight, cost).

<u>Values</u>: the terms, phrases, numbers, or other symbols which are available for assignment to elements for a given variable (e.g., red, 4 pounds, 50¢).

<u>Describers</u>: those values of variables which are assigned to particular elements.

<u>Observation/Measurement Procedures</u>: standard procedures or algorithms used to assign values of variables to particular elements (e.g., using a thermometer to measure the temperature of a liquid).

<u>Relational Rule</u>: rules or algorithms which specify describers for one variable given describers for another variable (e.g., $A = \pi r^2$, all the rectangular blocks are green).

<u>Correspondence Rules</u>: sets of rules used to relate one set of elements to another set of elements (e.g., the letter <u>p</u> is pronounced /p/).

<u>Text Referenced Instructional Management Systems</u>

Tests and texts have traditionally been treated as independent units with given tests amenable to various texts and the outcomes of instruction with a given text assessable by various tests. It is possible however to produce tests referenced to a given text series. With the test directly coupled to the text a means is provided for determinig the extent to which specific outcomes are being attained by individual students after specified instruction. It is also possible to prepare supplementary practice materials referenced to each criterion measure for use where adequate proficiency is yet to be attained. This integrated sequence of "text-test-troubleshooting materials" constitutes a simple instructional management system, which SWRL for convenience has termed a Learning Mastery System (LMS).

A prime limitation in producing such systems is that current texts rarely have clear statements of instructional outcomes. This limitation has been met by inferring the measureable outcomes associated with a given text. Although simple in structure and use, an LMS significantly expands the information available to the teacher for instructional decisions. Each LMS provides:

- A means for student placement at the beginning of the school year
- Criterion-referenced measures on three to eight instructional outcomes ten to fifteen times during the year
- Additional practice materials for the outcomes which have continuity throughout the text
- Mid-year and end-of-year evaluation measures.

Multiple Matrix Sampling

The specific equations used in multiple matrix sampling provided by Lord (1960) and Lord and Novick (1968) have been procedurally adapted for implementation (Shoemaker, 1973) and applied to large scale group achievement assessment. Results to date indicate that parameters estimated through multiple matrix sampling and parameters obtained through testing all examinees on all items may be interpreted similarly. Parameters estimated through multiple matrix sampling may be contrasted with any predetermined standard defining the minimal level of acceptable achievement.

State-of-the-Art Statistical Analyses

The Laboratory's research and development activities require on-line access to large data files and considerable flexibility in manipulating, analyzing, and retrieving information. In addition to standard statistical and matrix manipulation utility packages, a capability has been developed for the continuous upgrading of an extensive library of computer program-building modules. This permits quick modification of computer program functions with a minimum of reprogramming for new procedures defined by staff.

MATTERS THAT ARE ON THE LEADING EDGE OF SWRL STATE-OF-THE-ART

The areas and activities described in this section include items that while not quite available as "shelf-items," will influence the "new" generation of SWRL instructional products.

Quality Assurance Systems

The release of a SWRL-developed instructional program requires demonstration that it has been used successfully to obatin prespecified levels of pupil performance. To provide a replicable means of insuring that the program continues to function at these levels, a set of procedures referred to as Quality Assurance (Hanson, 1972) has been developed. These procedures provide en-route information on various indicators of performance and pacing useful to teachers, principals, and district administrators. Teachers have benefited from Quality Assurance because it provides information helpful in planning and pacing instructional activities throughout the school year. Principals and district personnel find Quality Assurance helps keep them informed of the status of an instructional program in each class throughout the school year. Pupils also benefit because it provides teachers with the assistance needed to complete all instructional units and to achieve high performance on the major outcomes.

Integrated Instructional Information Systems

Text referenced instructional management systems assign the teacher total accountability for the attainment of instructional outcomes. While the teacher often passes the responsibility on to the students, and

occasionally other school personnel, the teacher at present is the sole manager of instruction.

The confounding of the teacher, instructional materials, and instructional decisions in assessing accountability fails to recognize that the teacher shares responsibility for the instructional progress of students with administrators at the school and district level, and with parents. It is possible to provide useful information to each of these groups. However, the mechanisms for doing this are sufficiently complex to require automation of analysis and reporting functions. This is the scope of the SWRL Instructional Management System (IMS).

The SWRL Instructional Management System operates in conjunction with a developed instructional system such as the SWRL/Ginn Kindergarten Program or with an application of the Text Referenced Management System. Utilizing a variety of communication modes for input and output, reports for each category of individuals are specially designed to aggregate and synthesize the information in a manner that is understandable and comprehensive, consistent with need-to-know requirements of teachers, principals, curriculum supervisors, district administrators, parents, students, and development personnel (McManus, 1973).

Program Fair Evaluation

In the SWRL context "program fair" simply indicates that all assessment procedures are systematically referenced to the particular objectives of the program and the stimulus content used in instruction related to the objectives. Shoemaker (1972) has reviewed the state-of-the-art in this area. These techniques provide fair appoximations for

"program fair" comparisons of instructional programs. The adequacy of the approximations can only be assessed after the techniques have been further exercised empirically.

The Architecture of Instructional Programs

The item form and the instructional specifications (IS) are useful tools in instructional product development, but they are neither necessary nor sufficient to initiate or advance a given product development effort. Sets of IS's make it possible to define "trees" at an intermediate level of complexity above the micro-level of behavioral objective "twigs" but below the macro-level of an architectural framework. The architectural framework of an instructional program converts the "jungle" of instruction into an orderly "forest" configuration.

Emulating established procedures in the architecture of physical structures, the architecture of instruction can be conducted in stages of schematic specifications, through preliminary specifications, to working specifications. Instructional architecture subsumes the planning of "skills" and "content" conventionally considered in test design. Statements of instructional architecture are as yet few and far between. Examples of preliminary specifications can be found in Quellmalz (1973). An example of working specifications can be found in a SWRL (1972) document prepared by Baker, drawing upon various previous SWRL papers.

Instructional Development Control and Monitoring System (IDCMS)

IDCMS represents an integrated hardware configuration presently being installed within the Laboratory facility. It represents a powerful tool for increasing the sophistication of educational research and development activities. Computer applications to behavior research

have typically been restricted to statistical analyses of data collected off-line. This type of requirement can be handled by standard statistical and matrix manipulation utility packages. Although such a capability is important, Laboratory product design requirements include studies of real time inter-actions between subject and equipment. Exploitation of IDCMS capabilities will permit on-line experimentation involving complex event sequences, variable media utilization, and real time test contingencies. Figure 1 includes a block diagram of the IDCMS configuration.

MATTERS THAT ARE BEYOND SWRL STATE-OF-THE-ART

In an R&D context an instructional product that completes all stages of the development cycle is considered final; the "now" generation of the product unashamedly represents the best that can presently be delivered. However, long before the "now" product has gone to market the outcomes of progammatic R&D activity provide the scientific and technological bases for the "new" generation. Listed below are some items that were they now even "leading edge," the description of "new" generation products would likely be dramatically different. Yet, until they are classed as available shelf items the "new" generation of instructional products cannot be expected to reflect them.

1. Instructional data base structures in fields other than mathematics and reading.
2. Systemic structures in the social domain.
3. Cost-feasible automated interactive instructional hardware/software systems.

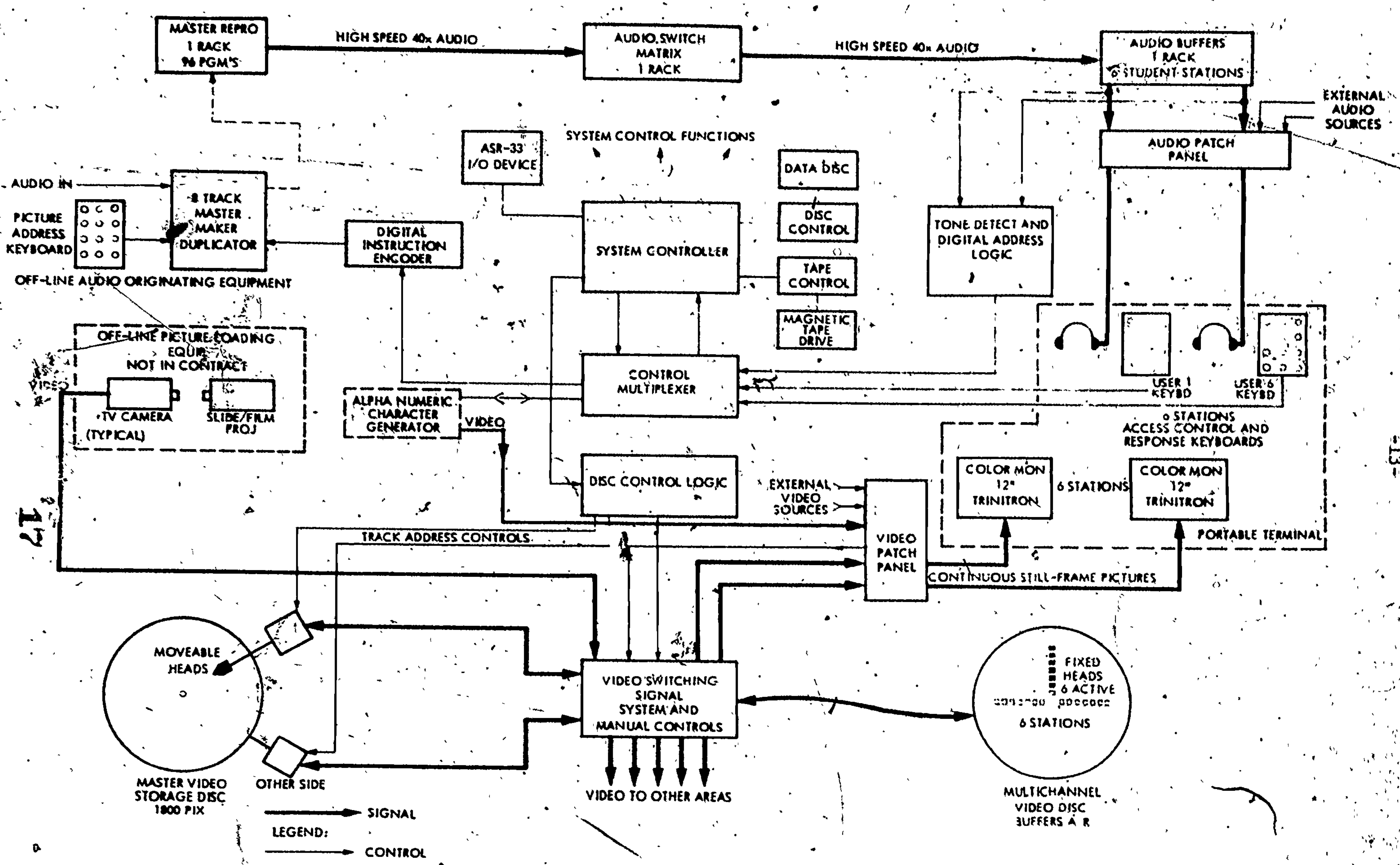

Figure 1

INSTRUCTIONAL DEVELOPMENT
CONTROL AND MONITORING
BLOCK DIAGRAM

4. Algorithms for prespecified instructional decision contingencies.

5. Quality control systems for aspects of performance other than qualitative attainment, time, and cost.

That the development of these items will involve measurement considerations is clear. These considerations move far from such classical topics as validity, reliability, item analysis, norming, and other traditional tools of psychometric theory and practice. It is well to have these tools in the instructional development kit, but more sophisticated tools are clearly needed.

REFERENCES

Bruner, J. S. The process of education. Cambridge, Mass.: Harvard University Press, 1960.

Baker, R. L. & Schutz, R. E., eds. Instructional product development. New York: Van Nostrand Reinhold Company, 1971.

Ebel, R. L. Criterion-referenced measurements: limitations. School Review, 1971, 79, 282-288.

Hanson, R. A. The contribution of quality assurance procedures to laboratory and user evaluation of educational programs. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April, 1972.

Hively, W. Defining criterion behavior for programmed instruction in elementary mathematics. Cambridge, Mass.: Harvard University, Committee on Programmed Instruction, 1963.

Lennon, R. T. Assumptions underlying the use of content validity. Educational and Psychological Measurement, 1956, 10, 294-304.

Lord, F. M. Use of the true-score theory to predict moments of univariate and bivariate observed-score distributions. Psychometrika, 1960, 25, 325-342.

Lord, F. M. & Novick, M. R. Statistical theories of mental test scores Reading, Mass.: Addison-Wesley, 1968.

McManus, J. F. The scope of IMS Version 3. In McManus, J. F., ed. Design of an instructional management system: SWRL working papers, 1972. SWRL Educational Research and Development, Los Alamitos, California, 1973.

Quellmalz, E. Structure of a fine arts instructional program. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, February, 1973.

SWRL. Technical specifications: English language and concepts for Spanish speaking children. SWRL Educational Research and Development, Los Alamitos, California, 1972.

Schutz, R. E. Criterion referenced testing. Paper presented at the annual convention of the American Educational Research Association, Chicago, April, 1972.

Shoemaker, D. M. Evaluating the effectiveness of competing instructional programs. Educational Researcher, 1972, 1, 5-8.

Shoemaker, D. M. Principles and procedures of multiple matrix sampling. Cambridge, Mass.: Ballinger Publishing Co., 1973.

Smith, E. L. Procedures for generating candidates for learning hierarchies. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, April, 1972.

Sullivan, H. J., Baker, R. L., & Schutz, R. E. Developing instructional specifications. In Baker, R. L. and Schutz, R. E., eds. Instructional product development. New York: Van Nostrand Reinhold Company, 1971, pp 65-104.